North Carolina's Proposal to Pilot the Use of a Growth Model for AYP Purposes in 2005-06

North Carolina has used growth as a part of its state accountability system, the ABCs of Public Education, since the 1996-97 school year. The growth standard was set based on observable changes in student performance from one year to the next. The acceptable standard was the average change across the state for all students including those who were already proficient. Using this system, the ABCs, for nine years, the achievement gap among ethnic groups appeared to be narrowing. After eight years using the same growth formulas as originally developed, North Carolina went through the tremendous task of reviewing the results of these formulas to determine areas where the process could be improved. The outcome of this review is a process in which individual growth targets are based on a student's prior achievement without regard to any demographic factors.

Using the processes already in place, North Carolina is proposing to use a modified form of its ABCs growth model, to add an additional layer to decrease the likelihood of falsely identifying schools as being in need of improvement that are providing quality education for their students. The proposal is based on recommendations from numerous education stakeholders in North Carolina including the Compliance Commission for Accountability, the Title I Committee of Practitioners, the Department of Public Instruction (DPI) Accountability Advisory Committee, and feedback from local school superintendents, principals, and central office test directors. The recommendations also were approved by the State Board of Education at its meeting on February 2, 2006.

After all other statistical methods and safe harbor have been applied to a school's proficiency targets; a growth trajectory would be calculated for all non-proficient students. This growth trajectory, should the student meet the trajectory's intermediate targets, would classify a student as performing "proficient" within four years in the tested grades. These targets are set based on initial status derived from the first test in the student record and project out to the grade-level test three school years later. Students who are on their trajectory in the current year would then be added to the proficient students for purposes of calculating proficiency against the Annual Measurable Objectives (AMO).

The details of these calculations and the logic behind the growth trajectories are discussed in Appendix A for growth.

The rest of this narrative will be divided into the categories provided in the Peer Review Guidance for simplicity.

1.1   How does the State accountability model hold schools accountable for universal proficiency by 2013-14?

The state will maintain its current AMO intermediate steps rising to the target of universal proficiency in the 2013-14 school year. These targets apply both for schools and LEAs. The growth trajectory included with proficiency will ensure that by 2014, all students will either be proficient or on their trajectory toward proficiency. In this calculation, students on trajectory will be added to proficient students to be compared against the AMOs. In this way, the model proposed is a Status, Safe Harbor, Growth model.

Based on reflective use of the proposed rules, an additional 40 schools would have met AYP using the proposed model (of the 932 schools that missed under the previous model in the 2004-05 school year). However, this is an overestimate of the effect because we have not had the time to calculate the number of years that a student has been in a tested grade in North Carolina and therefore we allowed all students to be eligible for the growth trajectory for these initial simulations.

The trajectory calculation would apply for 3rd grade (using the pretest administered at the start of 3rd grade as a baseline) through 8th grade (with 10th grade proficiency on the tests used for AYP at the 10th grade as the target).

1.2   Has the State proposed technically and educationally sound criteria for "growth targets" for schools and subgroups?

Based solely on student performance and ignoring demographic factors, North Carolina proposes that its growth trajectories will bring students into proficiency within four years of entering the tested grades in North Carolina. There are a variety of reasons for choosing this structure. The data are available to support the use of this model, and this model carries a student into proficiency within a reasonable length of time (although somewhat accelerated compared to student performance observed over the past nine years of NC data). Using the ABCs growth standard, only those students just barely below the proficient level would normally be expected to become proficient within a few years. More importantly, it was doubtful that any student who was in the lowest category of performance could become proficient within three school years of entering the tested grades.

Although described in more detail in Appendix A, the growth trajectory is built based on students' previous test scores compared to proficiency at a later point in time. For a non-proficient 3rd grade student (as an example), their pretest score from the start of third grade will be converted onto a common scale. The numeric difference between their 3rd grade pretest score and the common scale score to be proficient at the end of 6th grade will be calculated. At the end of third grade, if their common score has closed the distance from 3rd grade pretest to proficient at the end of 6th grade by 25%, the student

would count in the calculations as being on target. Similarly, the same student could count in 4th grade if the student's (second year) score has closed the distance by 50%. In 5th grade (third year), the distance would need to be closed by 75%. Finally in the student's 6th grade year (fourth year), the trajectory places the student at the proficient level. Note: this is a special case since there is a pretest at the beginning of 3rd grade.

These students on trajectory will be added to the number of students proficient in determining if the school or LEA has met its AMO for the group(s) to which the student is a member.

In the method above, OAI (other academic indicator) and participation targets are specifically excluded from the calculations. If a group misses one of these targets they cannot recover using growth calculations. The educational rational is straightforward. Participation is not a function of growth, schools either administer the appropriate assessments to their students or they do not. In the same way, not meeting the OAI for these groups does have an impact on the proficiency of the group and could encourage inappropriate practices. North Carolina also does not use a uniform averaging method for proficiency determinations.

To reiterate, the subgroup size and AMO targets are the same for each subgroup; there is no differentiation. In this way, our proposed methods still directly maintain and adhere to the original tenant of NCLB – closing achievement gaps between groups.

1.3   Has the State proposed a technically and educationally sound method of making annual judgments about school performance using growth?

The method proposed follows a process presented at national conferences/meetings and makes no adjustments for differences in student background characteristics. It is straightforward and easy to understand conceptually.

In North Carolina, proficiency AMO's were baselined using the method described in the NCLB legislation and subsequent regulations. Using stakeholder input, the decision was made to have three-year increases in proficiency goals on the way to universal (100%) proficiency. The growth proposal honors the intent of this method by aligning the targets for the growth component to the established proficiency goals. Each year, schools still will be expected to reach a proficiency rate at least equaling the proficiency goals stepping toward universal proficiency. Should the school not meet this target using a strict status model the following method is proposed. The number of students making the necessary gains to be on their individual trajectory to proficiency within four years in the tested grades in North Carolina, will be added to the number of students proficient to compare against the established proficiency AMO.

School and LEA AYP determinations would, under the proposed method, first follow all rules currently accepted under North Carolina's accountability plan. These measures include: comparison against the AMO directly, then the use of a 95% confidence interval, followed by the use of safe harbor provided the necessary data exist. The incorporation

of growth will be used as a final quality control on AYP determinations to ensure a decrease in false negatives. There will be no use of a confidence interval on the results of growth trajectory targets.

Although not used in determinations of AYP status, the growth of all students (both proficient and non-proficient) will be noted in the public AYP report.

1.4   Does the State proposed growth model include a relationship between consequences and rate of student growth consistent with Section 1116 of ESEA?

Using the proposed growth model, the series of consequential outcomes will remain unchanged from those currently accepted under the state accountability plan. In general terms, Title I schools who miss any targets (either participation or proficiency) in one subject two years in a row will enter improvement status. They will remain in that status, progressing in consequences each year that they miss any target in that subject until such a time that they make all targets (participation and proficiency) in that subject two consecutive years.

The intent of the proposed model is to decrease the number of schools falsely identified as being in need of improvement. These identifications are a drain on the limited resources available and dilute the effectiveness of interventions in the schools that are correctly identified as being in need of improvement.

2.1   Has the State proposed a technically and educationally sound method of depicting annual student growth in relation to growth targets?

The growth trajectory of students will be built using a method detailed in Appendix A. Briefly, to build the growth trajectory, we must be able to use a student's starting point (initial test score – can be thought of as a pretest score) and hopeful ending point (score for proficiency at some future point in time) and determine if the student's actual scores in the interim are at or above that trajectory. The student's initial score for most third graders is the pretest administered at the beginning of the third grade. For other grades, the end-of-grade assessment from the previous year is used as the pretest. The target is the score on the state's growth scale that is equivalent to the passing score on the test administered in the fourth year in the tested grades in North Carolina. For each year, the trajectory target is a 25% decrease (more if a student starts at a later grade level) in the difference from the pretest score on the state's growth scale to the score necessary to be proficient on the test in the fourth year in the tested grades in North Carolina. Using this method, a student's position on a trajectory path could be determined and documented as on- or off-trajectory in any given year.

For students who are lacking the necessary pretest scores, or certain students who use alternate assessments that are not on the growth scale, their participation is limited to their absolute status. It is important to note that proficient students are not included in growth trajectories for AYP purposes. However, North Carolina will use its ABCs growth formulas for proficient students to determine the percentages of all students

meeting growth expectations at a subgroup and school level. A student who scores proficient is weighted the same as a student who is on target. In this way, the growth of high-performing students does not compensate for the lack of growth among other students.

As the state replaces tests with new editions, the conversion between the state's growth scale and the required performance for proficiency in a certain grade will be determined. As a part of this process, the state standard for trajectory will be reviewed and modified to meet the needs introduced by the change in test edition.

3.1 Has the State proposed a technically and educationally sound method of holding schools accountable for student growth separately in reading/language arts and mathematics?

As described in detail in Appendix A, in each case, only scores from reading are used to calculate reading results and the same for mathematics. Student performance in reading is completely insulated from student performance in mathematics. Thus results remain separate and clearly delineated between reading and mathematics.

In schools with high mobility, the trajectory approach still includes all full academic year (FAY) students in the calculations because even students without the required pretest scores will be included based on their proficiency status.

4.1 Does the State's growth model proposal address the inclusion of all students, subgroups and schools appropriately?

The school as a whole and the subgroups are required to meet the 95% participation rate. If the school does not meet this target, they have not met AYP and growth cannot compensate for this issue.

Any full academic year student who participates in a valid test administration will be included in the growth calculations either on the basis of proficiency (when baseline scores are not available) or in the trajectory calculation. No modification is made to the minimum N size of 40 for a subgroup.

5.1   Has the State designed and implemented a statewide assessment system that measures all students annually in grades 3-8 and one high school grade in reading/language arts and mathematics in accordance with NCLB requirements for 2005-06, and have the annual assessments been in place since the 2004-05 school year?

The state has had a statewide assessment system in place since the 1992-93 school year including both reading and mathematics tests in grades 3-8 and high school assessments since the mid-1980's. The mathematics end of grade assessments for grades 3-8 will be a new edition in the 2005-06 school year. This will not be an issue with trajectory growth since the tests will have a conversion to the state growth scale and they will be equated in terms of proficiency cut scores to the previous edition.

5.2   How will the State report individual student growth to parents?

No plans are in place currently to report individual student growth scores to parents. The reason for doing so is based on the state's decision to not report individual results to parents in any terms other than status on the developmental scale, achievement level and achievement level descriptors.

The state, however, will report AYP group level growth results as part of the public AYP reporting. Although this percent of students meeting their growth targets will not be used in making AYP determinations (since it will include both proficient and non-proficient students), it will give parents and other stakeholders insight into the functioning and effectiveness of the schools in NC. It will also provide data and insights for researchers looking at the possible impact of using growth models for accountability purposes.

5.3   Does the Statewide assessment system produce comparable information on each student as he/she moves from one grade level to the next?

Yes, the state assessment system has been developed to provide both vertical scaling and the capacity to calculate growth. As part of the state's current accountability system individual student growth is calculated using a scale specifically for growth calculations.

5.4   Is the Statewide assessment system stable in its design?

Yes. The state's assessment system was one of the first to undergo the Peer Review process and we anticipate the USED requests for additional information from the review to be cleared during the same time as the review of this proposal.

6.1   Has the State designed and implemented a technically and educationally sound system for accurately matching student data from one year to the next?

Over the previous nine years, the state has used a system of drawing data back from LEAs for the purposes of growth calculations. This system is aligned with state policies and has served well for the purposes of calculating growth. Each year cross referencing is done at the state level to identify mismatched test scores and the LEAs receive feedback to correct any errors identified. In this way, the school that receives the cumulative records for a transfer student uses them in their student information management system. Audits are run locally by the school and LEA prior to submission to the SEA for verification and error flagging.

Match rates across years are strong and detailed in Appendix B. Match rates within a year (student records that have the necessary pretest scores) yields better than 93% of students have the necessary pretest scores. The percentages of LEP students and Hispanic students who have pretest scores are lower due to the number of students in that category who either move into the state or have the first year exemption from being included in the assessment system in reading. Of the student pretest scores, 97.4% of the scores have been validated as matching for the student the previous year, the rest are either the result of a misadministration or have been sent back to the school for reconciliation. In this analysis, no general trend exists in the mismatch.

Again it is important to note that a student who does not have the pretest scores to calculate a growth trajectory will still be included in the AYP determinations solely on the basis of proficiency status.

6.2   Does the State data infrastructure have the capacity to implement the proposed growth model?

The mainstay of North Carolina's accountability system is its growth model. This model has been in place since the 1996-97 school year and has functioned well. The data have supported the model and there has not been any observable data issue hindering the functioning of the growth model.

7.1   Has the State designed and implemented a statewide accountability system that incorporates the rate of participation as one of the criteria?

Yes, in the proposal, the state holds the school and LEA to at least a 95% participation rate and trajectory growth cannot compensate for the participation rate.

7.2   Does the proposed State growth accountability model incorporate the additional academic indicator?

Yes, schools and LEAs are held to the other academic indicator (attendance for schools that do not graduate students and graduation rate for high schools) and growth cannot compensate.

Appendix A

A Four-Year Proficiency Growth Trajectory Model

As part of the state accountability system, North Carolina uses both End-of-Grade (EOG) tests scored on a developmental scale and End-of-Course (EOC) tests scored on a discrete scale. During the 2004-05 school year, a legislatively mandated review of the growth standards was completed. The review prompted a change to a Standardized Scale Approach (SSA) to growth which uses the normative distribution of student performance in the standard setting year of any test edition as a common basis to build a scale. This approach is useful for measuring the growth in student performance from one year to the next and also adapts well to the changes in curriculum and subsequent changes in test editions.

The SSA system uses a time-locked modified z-scale. This system has been endorsed by the educational stakeholders and the state's Technical Advisory Committee (TAC). The scale is termed a "change scale" or "c-scale." Thus, the c-scale cut score for proficiency on any given test edition at an individual grade level remains constant for the life of the scale and test edition regardless of the changes in the distribution of test scores that might occur as schools change their instructional methods. The state means and standard deviations from the standard setting year are used indefinitely for any given test.

The 2005-06 school year is the standard setting year for the Mathematics EOG tests at grades 3-8. It is anticipated that the 2007-08 school year may be the standard setting year for the new Reading EOG tests in grades 3-8. It is also important for the discussion of proficiency to note the equating study that sets the achievement level cut scores is performed at the same time the c-scale is built.

The following tables contain the means and standard deviations used to convert from the developmental or discrete scales to the c-scale for growth purposes.

**Table 1. Standard Setting Years, Means and Standard Deviations
for End-of-Grade (EOG) C-Scale Computation**

EOG	Standard Setting Year	Scale Score Mean	Standard Deviation	C-Scale Score Needed to Reach Proficient
Reading (2nd Edition)				
Grade 3 Pretest	2003	238.7	9.94	-0.57
Grade 3	2003	247.9	9.06	-0.87
Grade 4	2003	252.3	8.68	-0.96
Grade 5	2003	256.9	8.03	-1.23
Grade 6	2003	258.7	8.55	-0.78
Grade 7	2003	261.1	9.06	-1.00
Grade 8	2003	263.9	9.05	-1.09
Mathematics (2nd Edition)				
Grade 3 Pretest	2001	236.1	8.10	-0.75
Grade 3	2001	250.6	7.75	-0.59
Grade 4	2001	255.8	8.32	-1.06
Grade 5	2001	260.0	9.62	-1.04
Grade 6	2001	263.2	9.91	-0.92
Grade 7	2001	267.1	10.63	-0.86
Grade 8	2001	270.0	10.95	-0.82

All values are rounded to either one or two decimal places for the table. Full precision was used for actual calculations.

Table 2. Standard Setting Years, Means and Standard Deviations for End-of-Course (EOC) C-Scale Computation

EOC	Standard Setting Year	Scale Score Mean	Standard Deviation	C-Scale Score Needed to Reach Proficient
Algebra I	1994	55.1	9.12	
	2001	61.1	9.31	-0.66
English I	1995	53.1	8.94	
	2003	57.7	7.63	-0.75

All values are rounded to one or two decimal places for the table. Full precision was used for actual calculations.

The state's normal growth expectation under the ABCs accountability program is the following:

$CS_{c\text{-}scale} = (0.82 \times PA_{c\text{-}scale})$

Where:
- CS = current score
- PA = previous assessment score

Using the logic and building on the technical background provided at:

http://www.ncpublicschools.org/accountability/reporting/growthformulas

The trajectory is built based on the student's performance either the previous year, or on the 3rd grade pretest, whichever is appropriate to the grade in which the student first enters the state. Therefore, the following table illustrates the basis for prediction, the targeted test for proficiency, the years of trajectory, and the percent of difference between baseline performance and proficiency expected by the trajectory based on the year the student is first enrolled in the state in a tested grade.

**Table 3. Grades And Tests Used For Trajectory Growth And
The Percent Of Closing Expected Per Year**

Grade Of First Enrollment	Test Used As The Basis For Prediction	Test Used As Target For Proficiency	Years To Proficiency	Percent Of Difference Closed Per Step	Steps To Proficiency
3	3rd grade pretest	6th grade EOG	4	25%	4
4	4th grade EOG	7th grade EOG	4	33%	3
5	5th grade EOG	8th grade EOG	4	33%	3
6	6th grade EOG	Algebra I or English I EOC	4	33%	3
7	7th grade EOG	Algebra I or English I EOC	4	50%	2
8	8th grade EOG	Algebra I or English I EOC	3	100%	1

The trajectories are built individually by student and separately for reading or mathematics. Therefore, a student will have a trajectory based on their baseline mathematics score and the proficiency cut score for mathematics separate from reading. In the upper grades, Algebra I is the AYP assessment for 10th grade students and is the trajectory target for math while English I is the trajectory target for reading/language arts.

The following table displays the performance expected of students to be counted as on trajectory for inclusion in the proposed method of comparing school performance to AMO targets.

**Table 4. The Amount Of Improvement In Terms Of Decrease In The Distance
Between Baseline Performance And Proficiency In The Target Grade**

For A Student Who Enters In 3rd Grade and Has a Grade 3 Pretest

Year In State-Tested Grade	Decrease From Baseline Assessment In Performance Discrepancy
1	25% Of Original Gap
2	50% Of Original Gap
3	75% Of Original Gap
4 or more	Student Must Be Proficient

For A Student Who Enters In 4th, 5th, or 6th Grade

Year In State-Tested Grade	Decrease From Baseline Assessment In Performance Discrepancy
1	Baseline, Not On Trajectory
2	33% Of Original Gap
3	66% Of Original Gap
4 or more	Student Must Be Proficient

Therefore, if a subgroup has met its 95% participation target but has not met its proficiency target, and the subgroup has met its other academic indicator, the process would be:

1) First identify if the student has been in membership the full academic year and is both tested and not proficient.
2) These three conditions being met, the number of years the student has been in the state will be determined using the historic files from the state's accountability system.
3) If the student has been in the state (in a tested grade) for four years or more, the student will remain non-proficient for comparison to the annual measurable objectives (AMO). If the student has been in the state public schools three years or less, the correct baseline score will be located (using the table above).
4) The student's performance on the baseline assessment in the subject of interest will be converted to the c-scale.
5) Based on the student's baseline score and proficiency in the target year, a difference will be calculated.
6) The decrease in the difference will be compared against Table 4 above based on the number of years in the tested grades in North Carolina.
7) If the student's performance on the current assessment is equal to or better than the minimum from the previous step, include the student in the percent proficient calculation to compare against the state's AMO's.

An example follows:

A student enters North Carolina in the 4th grade. The student scores below proficient in the current school year in reading. This child's known test scores are listed below.

Grade	3 EOG	4 EOG	5 EOG
Developmental Score	Not in NC	229	241
C-scale score		-2.68	-1.98

Since the student's first full year in the state is the fourth grade year, the student will need to be on trajectory to be proficient by the end of the seventh grade and thus on the seventh grade EOG for reading. The developmental score for seventh grade reading equivalent to proficient is 252. The associated c-scale score is -1.00.

Since the student was not in the state for the third grade test, the fourth grade EOG score will be used as the baseline. The difference between the baseline and proficient on the seventh grade test in terms of c-scale scores is 1.68 (difference between 2.68 and 1.00). For the current year (fifth grade, the second year in the state), the student must perform well enough on the test to have 33% less difference between the c-scale score for proficiency and his baseline (4th grade EOG) c-scale score (divide 1.68 by 3 = 0.56).

For this to be true, the child would need to score at least -2.12 (difference between 2.68 and 0.56). The child's actual c-scale score is -1.98 which means the child met the

standard to be deemed on trajectory for the current year and thus will be included in the percent of students on trajectory or proficient for comparison to the AMO for the school as a whole and any subgroups the child may be a part of.

Appendix B

Percent Of Students With Pretest Scores And Quality Of This Data

Subgroup	Pct of Students with Pretest Reading 2004-05	Pct of Students with Pretest Math 2004-05
American-Indian	96	96
Asian	90	90
Black	94	94
Free and Reduced Lunch	93	93
Hispanic	83	86
Limited English	73	78
Multiracial	91	91
Student with Disabilities	94	94
White	94	94
Total	93	93

Note: This analysis includes the students in grades 3 to 8 used in the ABCs growth calculations for 2004-05. All percentages rounded to whole numbers.

Quality of Pretest Scores
In the standard review of data submitted to DPI, 97.4% of the test scores reported for students this year, for tests taken last year, matched the scores from last year. Part of the discrepancy is based on the way misadministered tests are recorded (they appeared twice, once for the misadministration and once for the valid administration and this automatically reports as a mismatch). The other discrepancies are returned to the LEAs for reconciliation prior to a mid-year quality assurance data collection.